Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 7438–7442, November 1987
Biochemistry

Conservation of ferritin heavy subunit gene structure: Implications for the regulation of ferritin gene expression

(translational control/iron/RNA secondary structure/rat)

MARY T. MURRAY*, KRISTIN WHITE*†, AND HAMISH N. MUNRO*‡§

*Department of Applied Biological Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; †School of Nutrition, Tufts University, Medford, MA 02155; and ‡U.S. Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, MA 02111

Contributed by Hamish N. Munro, July 10, 1987

ABSTRACT **Ferritin stores iron within a protein shell consisting of 24 subunits of two types, heavy (H) and light (L). According to Southern blotting, the rat genome contains four copies homologous to the H-subunit cDNA (H cDNA). To determine whether only one of these is expressed, H cDNAs isolated from rat liver and heart mRNAs were compared and found to share identical nucleotide sequences. Next, genomic clones for three of the four rat H-subunit loci were isolated. Two were classical processed pseudogenes, whereas the third contained an expressed gene. RNase intron mapping of this expressed gene generated the same exon protection pattern when total RNA from rat liver or heart was used, indicating that this gene accounts for most or all of the H-subunit mRNAs (H mRNAs) in these tissues. Comparison of the expressed rat H-subunit gene (H gene) structure with published sequences for other species displays considerable conservation. The coding sequence of the rat H gene predicts 95% similarity to the human amino acid sequence, thus being more highly conserved than the L-subunit sequence of these species. Near the cap region of the 5′ untranslated region, the rat H mRNA displays a 28-nucleotide sequence that is almost totally conserved in the corresponding region of the human, bullfrog, and chicken H mRNA and is also faithfully represented in the rat and human L-subunit mRNAs (L mRNAs), thus making this sequence a prime candidate for involvement in the known translational regulation of both subunits by iron. In the 5′ flanking region, partially conserved sequences common to H gene and L-subunit gene (L gene) of the rat may be involved in transcriptional regulation by iron, whereas those conserved only in the H gene of man and the rat imply that other factors may independently control H-subunit regulation.**

In most mammalian tissues, the iron-storage protein ferritin consists of a shell of M_r 450,000 made up of 24 subunits of two types, heavy (H) and light (L) (1). The proportion of each subunit varies from predominantly L in liver and spleen to predominantly H in heart (2). On analysis of liver and heart ferritins by electrofocusing, a series of isoferritins separates by charge according to the proportions of H and L subunits in their shells. However, heart and liver isoferritins with the same surface charge (pI) display markedly different proportions of H and L subunits (3), suggesting charge differences between the H and/or L subunits of heart and liver due either to variations in amino acid sequences between the tissue ferritin subunits or to posttranslational modifications.

The possibility of tissue-specific ferritin genes was increased by finding multiple gene copies of rat L (4) and human H (5) subunits by Southern blotting. However, the isolation of genomic copies encoding the human H subunit (6, 7) and human and rat L subunits (8, 9) has demonstrated the presence of several processed pseudogenes. In addition, isolation of H-subunit cDNAs (H cDNAs) from human liver, lymphocytes, HeLa cells, and endothelial cells yielded identical sequences (6), and only one of the numerous gene copies of the rat L subunit appears to be expressed (9). Accordingly, in order to complete the picture for the rat ferritin genes, we have examined the gene copies for the rat H subunit and report that only one appears to be expressed.

Finally, we have evaluated the structure of the expressed rat H-subunit gene (H gene) and its cDNA for potential regulatory features in the sequence. First, iron administration causes preferential transcription of the L-subunit gene (L gene) (10), whereas cell differentiation appears to favor accumulation of H-subunit mRNA (H mRNA) (11). Second, iron also induces synthesis of both H and L subunits by causing latent H mRNAs and L-subunit mRNAs (L mRNAs) present in the cytosol to become polyribosome-associated and become translationally active (12, 13). This suggests a common iron-sensitive signal on both H and L mRNAs.

MATERIALS AND METHODS

Isolation of Ferritin H cDNAs and Genomic Clones. (*i*) Using the procedure of Gubler and Hoffman (14), double-stranded cDNA was synthesized from poly(A)$^+$ RNA isolated from the hearts of young male Sprague–Dawley rats (15). A rat heart cDNA plasmid library of ≈30,000 clones was constructed from cDNA selected on a sucrose gradient to be >600 base pairs (bp). This library was screened on nylon filters (Hybond-N, Amersham; or Biotrans, ICN) (16, 17) with a human liver H cDNA (5), ^{32}P-labeled by the random-primer method (18). (*ii*) Fischer rat liver λgt11 cDNA (19) and Fischer rat genomic λEMBL3B (20) libraries, generously provided by R. O. Hynes, were screened in a similar manner. From the genomic library three rat H genomic clones were isolated; two proved to consist of processed pseudogenes (*5-13* and *22-1*), whereas the third clone contained an expressed H gene (*23-4*).

Southern Blotting. Genomic Fischer rat liver DNA was isolated from nuclei lysed by the addition of sodium sarcosinate to 4%. Proteinase K was added and the viscous solution was gently rotated at room temperature overnight; this was followed by fractionation of the DNA on a step cesium chloride gradient. Southern blots (21) on nylon filters (Hybond-N, Amersham) were probed with random-primer-labeled rat liver H cDNA (see Fig. 2*A*). The filter was then stripped of probe and reprobed with an M13-generated probe (9) of the 5′ flanking region (−177 to −35) of the expressed rat H gene (see Fig. 2*B*). For the mixed genomic clone Southern blot (see Fig. 2*C*), equal amounts of each of the rat H λ phage DNAs isolated by the cesium chloride preparative method (22, 23) were combined, restricted with *Sal* I to

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: H subunit, ferritin heavy subunit; L subunit, ferritin light subunit; UTR, untranslated region.
§To whom reprint requests should be addressed.

release the insert from the λ arms, and then restricted with the appropriate restriction enzyme. The restriction fragments were separated by agarose gel electrophoresis, blotted onto nylon membrane, and probed with the random-primer ^{32}P-labeled rat liver H cDNA.

DNA Sequencing. Ferritin H genomic and cDNA M13 subclones were sequenced by the dideoxy method of Sanger *et al.* (24) using deoxynucleotide 5′-[[^{35}S]thio]triphosphates. A synthetic 16-base oligonucleotide primer (complementary to nucleotides 1133–1148 in Fig. 1) was used to determine the 3′ flanking sequence. All coding sequences and the 3′ untranslated region (UTR) were sequenced in both directions. Since the 5′ flanking region and UTR were (G+C)-rich, 7-deaza-2′-deoxyguanosine 5′-triphosphate (Boehringer Mannheim) was used in sequencing these regions. The 5′ flanking region and UTR were sequenced by the dideoxy method in the 5′ to 3′ direction and verified by Maxam–Gilbert (25) sequencing from the *Bam*HI site 5′ to the TATAA box.

RESULTS AND DISCUSSION

Sequence of Rat Ferritin Heart and Liver H cDNA Clones. Although neither was full length, the nucleotide sequences of the two cDNA clones were identical with one another and with that of the expressed H gene (Fig. 1). The 5′ UTRs of the liver and heart H cDNA were incomplete, the liver having 90 bases and the heart only 3 bases, compared with the corresponding sequence of the expressed H gene, which indicates 168 bases in the intact 5′ UTR (Fig. 1). The reading frame of 543 bases and the 152 bases of 3′ untranslated sequence of the H gene were faithfully repeated in the two tissue mRNAs, indicating that they were identical transcripts of the expressed H gene. From the derived amino acid sequence it can be calculated that the rat H subunit has a M_r of 21,004, whereas the rat L subunit has a M_r of 20,700 (26), the molecular weight being reflected in the close mobility of the H and L polypeptides in denaturing gel electrophoresis (3, 4).

Characterization of Genomic H Clones. Three distinct rat H genomic clones were isolated and characterized by restriction mapping, RNase intron mapping (27), and nucleotide sequencing. Two of these genomic clones proved to contain processed H pseudogenes (*5-13* and *22-1*), whereas the third was an expressed H gene (*23-4*); these are identified in the Southern blot shown in Fig. 2*C*. Genomic clones *5-13* and *22-1* are distinct processed pseudogenes; they each contain

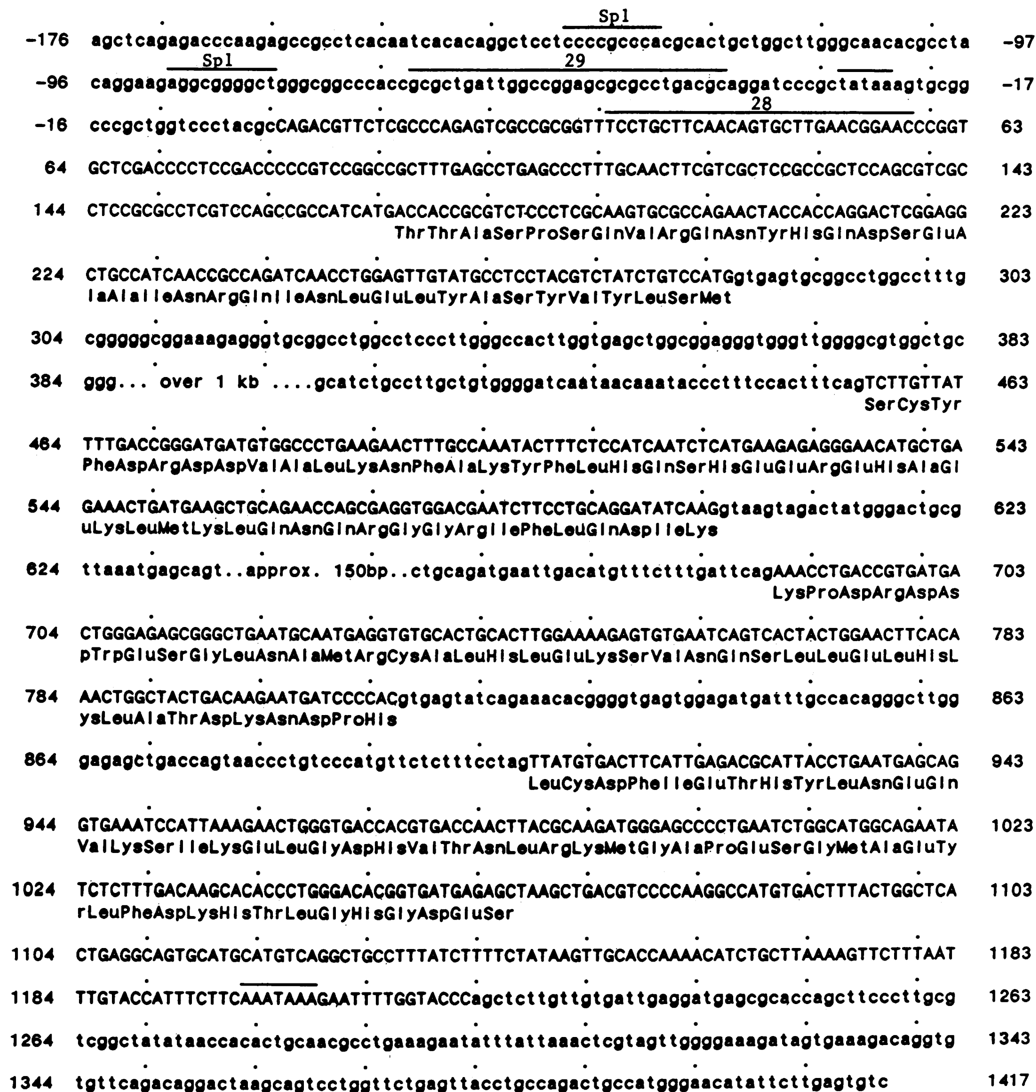

FIG. 1. Nucleotide and derived amino acid sequence of the expressed rat ferritin H gene *23-4* with the exons shown in uppercase letters. The nucleotides of the flanking regions and introns are shown in lowercase letters. In the 5′ flanking region the sequences identified by overbars are two Sp1 consensus binding sequences, a 29-bp sequence conserved between rat and human H genes, and the TATAA box. In the first exon, the indicated 28-bp sequence is conserved in the 5′ UTRs of human and rat H and L, chicken H, and bullfrog H. The poly(A) consensus sequence is indicated by an overbar in the 3′ noncoding region.

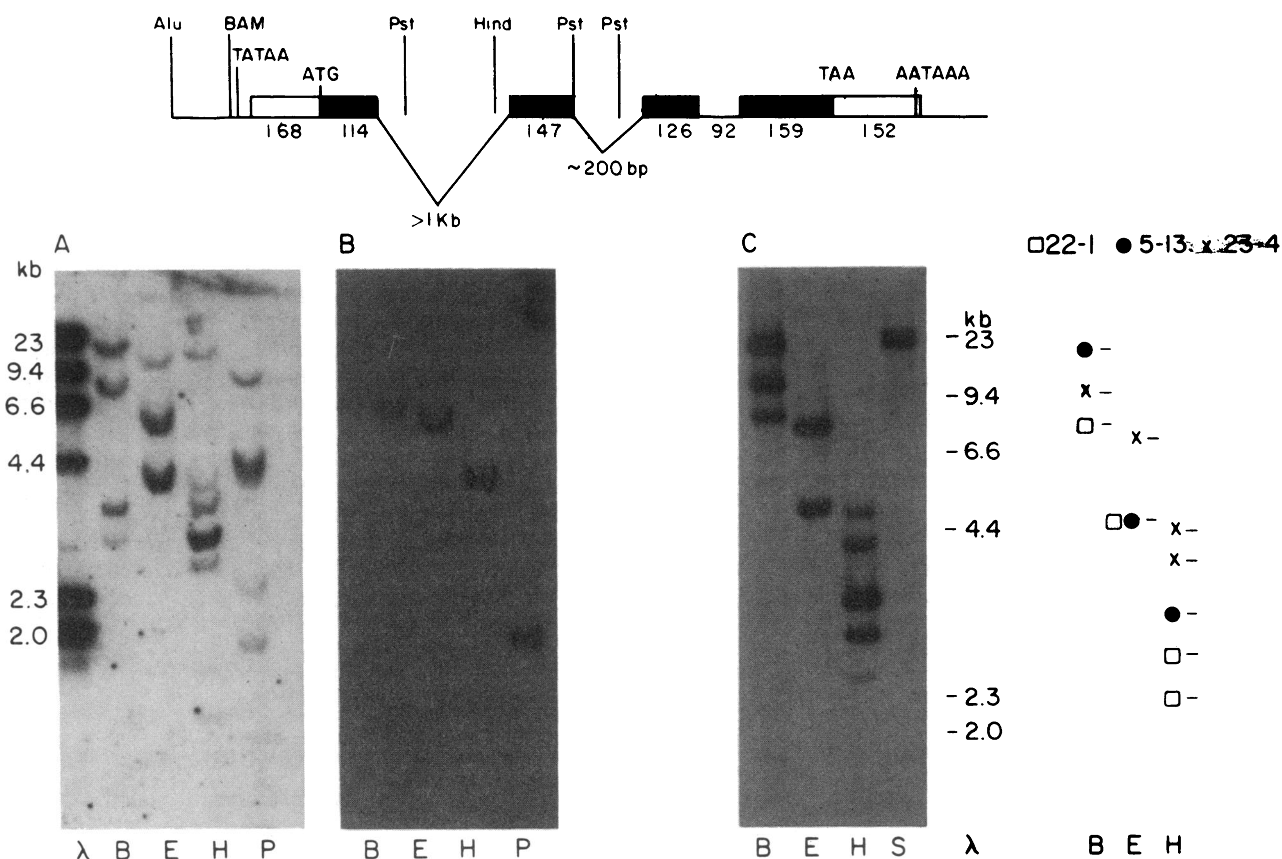

FIG. 2. Comparison of genomic rat ferritin H Southern blot to isolated H genomic clones and identification of a single genomic copy of the promoter region of genomic clone *23-4*. (*A* and *B*) The same Fischer rat genomic Southern blot probed with the rat liver H cDNA in *A* and reprobed in *B* with the 5′ flanking region (*Alu* to *Bam*HI) of genomic clone *23-4* (illustrated at the top). (Note that, in the *Bam*HI digest shown in *B*, the probe identifies a fragment not present in *A* because *Bam*HI separates this probe from the first exon.) (*C*) Southern blot of the isolated rat H genomic clones mixed together and the blot probed with the rat liver H cDNA. kb, Kilobases. The panel to the right of *C* identifies the positions in *C* of H genomic clones *5-13* and *22-1* containing processed pseudogenes and the position of the expressed H gene *23-4*. The restriction enzymes are indicated as follows: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst* I; S, *Sal* I. λ represents λ *Hin*dIII standards. In *C*, genomic clone *22-1* digested with *Bam*HI migrated faster than in *A* as a result of cloning into the λ vector.

the H processed transcript followed by poly(A) tails and are flanked by direct repeats. The direct repeats for *5-13* are 5′ GAAATCCTGGGTAAA.....GAAATCCTGGGAAA 3′ and for *22-1* are 5′ GAATGGATCAGTTTG.....GAAATGGATCAGTTTG 3′.

Using the expressed H gene (*23-4*) for RNase intron mapping, the same exon protection pattern was generated with either total heart or liver mRNA (data not shown), indicating that the H mRNAs of heart and liver are identical. This provides additional evidence that the same H gene transcripts are produced in liver and heart cells and that no additional H genes are expressed in detectable amounts. Similarly, we have determined that the nucleotide sequence of a rat heart L cDNA is identical to rat liver cDNA and that intron mapping of the rat L gene (9) using either rat liver or heart RNA generates the same protection pattern (unpublished observations). The above findings of the same H and L mRNA expressed in rat heart and liver do not support tissue-specific subunits but, instead, suggest differential posttranslational modification between heart and liver ferritin as the probable origin of tissue ferritin heterogeneity.

To explore whether the number of H homologous bands on Southern blotting can be accounted for by the three above genomic clones, a Fischer rat genomic Southern blot probed with the isolated rat liver H cDNA (Fig. 2*A*) was compared to a Southern blot of the genomic clones combined following their isolation (Fig. 2*C*). In the genomic Southern blot (Fig. 2*A*) the H cDNA hybridizes to four *Bam*HI, three *Eco*RI, and five *Hin*dIII fragments. When the isolated genomic clones were probed with H cDNA (Fig. 2*C*), all but one of the rat H genomic bands generated with *Eco*RI or *Hin*dIII (Fig. 2*A*) are accounted for by the isolated genomic clones. The H genomic bands that were not accounted for by the genomic clones in Fig. 2*C* reside on 10- to 20-kb *Eco*RI or *Hin*dIII fragments and contain an internal *Bam*HI site, thus accounting for two *Bam*HI fragments (determined by Southern blots probed with regions of the H cDNA; not shown). It is coincidental that the two processed pseudogenes (*5-13* and *22-1*) comigrate upon genomic *Eco*RI digestion and therefore are represented by a single hybridization band after this restriction. Therefore, all but one of the genomic bands can be accounted for by the isolated genomic clones. The evidence presented here and studies of human H gene expression (6) suggest there is a single H mRNA expressed in various tissues. If a second H gene exists at this unresolved locus, either it is not expressed in liver or heart or it represents only a small fraction of the total H mRNA in these tissues.

Characteristics of the Expressed Rat Ferritin H Gene. The third isolated gene sequence, *23-4* (Fig. 1), contained 177 bp of 5′ flanking region, 168 bp of 5′ untranslated, 153 bp of 3′ untranslated, and 196 bp of 3′ flanking sequence, as well as the coding region with its introns. The transcriptional initiation site was inferred by homology to the human H gene (6) and the rat H processed pseudogenes *5-13* and *22-1*. Gene *23-4* was identified as an expressed H gene because it is completely homologous to the H cDNAs of liver and heart

and its 5′ flanking region is partially homologous to the same region of the human H gene that has been expressed in mouse cells (7).

The coding region of the rat H gene is 90% homologous to the human H-coding sequence (5, 11, 28) and its derived amino acid sequence is 95% homologous. By comparison, the L-subunit amino acid sequence is only 85% conserved between rat and human (9). All of the residues of the H subunit postulated to be involved in intersubunit interactions and iron core formation and those contributing to the two types of channel traversing the apoferritin protein shell (28) are conserved between rat and human. The amino acid sequence derived from the H gene is in agreement with a partial amino acid sequence obtained by Sequenator for the rat liver H subunit from Met-166 to the C-terminal end (26), except that the gene predicts the terminal residue to be serine rather than glutamic acid.

The coding region of the expressed H gene is interrupted by three introns. The sizes of the first and second were estimated by restriction mapping to be ≈1500 bp for the first and ≈200 bp for the second, of which 1100 bp and 175 bp, respectively, were sequenced, whereas all 92 bp of the third intron were sequenced. Computer-assisted analysis using Chou–Fasman rules (29) predicted that the five helical regions determined by x-ray crystallography for horse spleen ferritin L subunit (30) are preserved in the H and L subunits of human and rat (28, 31). The positions of the three introns, which fall at hinge regions between α-helical domains, are conserved between the H (6, 7) and L genes (8, 9) consistent with their evolution from a common ancestral gene.

Evolutionary Conservation of UTR Sequences. The majority (>70%) of 5′ UTRs of a group of eukaryotic mRNAs surveyed have relatively short 5′ UTRs of 20–80 nucleotides (32), whereas that of the rat H mRNA is 168 bases long and the human is 209 bases in length; those of the L mRNAs are 203 bases for rat (9) and 198 bases for man (8). The 5′ UTRs of rat and human H mRNA tend to be (G+C)-rich (65% and 73%, respectively). The first 94 bases of the H 5′ UTR show 84% homology between rat and man, including a 28-base sequence (Fig. 1) that is highly conserved in the 5′ UTRs of human H (8), human and rat L (8, 9), chicken H (33), and bullfrog H (34). This conserved sequence may be significant for the translational control of the H and L messages, which are present in the cytosol in a translationally inactive form that becomes polyribosome-associated when iron is injected into rats (12) or added to cells in culture (13) or to bullfrog reticulocytes (35). In each of these species, the 28-base sequence can potentially fold back on itself, with the internal CAGUGC forming a loop (Fig. 3). The calculated free energy (36) for the stem-loop structures shown in Fig. 3 are −15.1 kcal/mol (1 kcal = 4.184 kJ) for the bullfrog and −40.3 kcal/mol for rat H. The sequences immediately 5′ and 3′ to the 28 bases, although not conserved between H and L mRNA, can potentially base-pair to form a larger stem for the loop according to computer analysis (37). The structure shown in Fig. 3 for the rat H has the same conformation as a predicted structure for the rat L, although the only residues in common are the conserved 28 bases. Deletion analysis of rat L 5′ UTR choline acetyltransferase chimeric constructs demonstrates that translational regulation by iron requires a region containing the conserved 28-base sequence (N. Aziz and H.N.M., unpublished).

Using sequences inserted in the 5′ UTR to form secondary structures, Kozak (38) has shown that translation can be blocked by structures with stabilities exceeding $\Delta G = -50$ kcal/mol, whereas the 40S ribosomal subunit can still migrate through less stable secondary structures with $\Delta G = -30$ kcal/mol. In agreement, purified ferritin mRNA is translated efficiently *in vitro* (4, 13, 35), indicating that the potential stem-loop structure by itself is not an impenetrable barrier without some additional repressor molecule. An iron-sensitive molecule blocking translation was originally postulated by Zahringer *et al.* (39) to account for the occurrence of latent ferritin mRNA in the messenger ribonucleoprotein fraction of liver and for its translation upon iron administration. Recently, RNA protection studies have provided evidence of a specific protein binding to the 28-base sequence, with the nature of its binding being altered following iron administration (E. Leibold and H.N.M., unpublished).

BULLFROG H RAT H

FIG. 3. Potential secondary structure in the 5′ UTR of bullfrog (bases 1–70) and rat H ferritin mRNA (bases 1–100). The conserved 28-base sequence is outlined in each potential structure.

Regarding the 3′ UTR, there is 85% homology between rat and human H mRNA over the terminal 130 bases of its 153 bases, whereas there is no significant homology of the rat H 3′ UTR to rat (26) and human L 3′ UTRs (31) or to the bullfrog 3′ UTR (34). Consequently, this conservation across the H 3′ UTRs of two species implies a function specific to the H mRNA, possibly related to the greater stability of H mRNA compared with L mRNA (unpublished data).

Conservation of the Promoter Region. The isolated expressed H gene is present as a single copy in the rat genome (Fig. 2*B*). A promoter region-specific probe generated from genomic clone *23-4* (−177 to −35, *Alu–Bam*HI) detects a single band on the Fischer rat genomic Southern blot (Fig. 2*B*) and identifies the expressed gene in Fig. 2*A*. The H locus that was not isolated as a genomic clone (compare Fig. 2*A* to 2*C*) does not contain sufficient homology to the promoter region of *23-4* for cross-hybridization (Fig. 2*B*), implying that if it does contain a H gene rather than a pseudogene, it has

control elements different from the isolated H gene, which is expressed in heart and liver.

The 5′ flanking regions of the rat expressed H gene (Fig. 1) and human H gene (6) contain a modest excess of G·C bases, 64% and 72%, respectively. Upstream of the TATAA box is a 29-bp sequence (−65 to −38), of which 27 bases are conserved in the human H-flanking region (6), and the last 6 are also found in the chicken H gene (33). In the rat (Fig. 1) and human H gene (6), this conserved sequence is preceded by an Sp1 consensus binding sequence (40). Conservation of the 29-bp sequence upstream from the TATAA box suggests that this is part of the H-gene promoter, which may also include the adjacent Sp1-binding sequence. This can be tested by DNA "footprinting" and by the response to iron of constructs from which these segments have been deleted.

Differences in the functions of the H and L subunits remain obscure. Differentiation of the human promyelocyte cell line HL60 to either macrophages or neutrophil leukocytes is associated with a complex pattern of H- and L-gene expression that favors accumulation of more H mRNA (11). The transcription of H and L genes is also differentially affected by iron administration, in this case L being favored over H transcription (unpublished experiments; ref. 10). Relevant to this action of iron are segments of 70% homology between the 5′ flanking region of the rat L gene at nucleotides −259 to −218 and the rat H gene at nucleotides −166 to −124 upstream from the cap site. These may participate in the complex pattern of ferritin subunit expression that varies according to the tissue and its iron storage status (3).

We are grateful to Dr. R. Cortese, European Molecular Biology Laboratory, Heidelberg, for providing the human liver H cDNA, and Dr. R. Hynes, Massachusetts Institute of Technology, for the rat liver cDNA and genomic libraries.

1. Theil, E. C. (1983) *Adv. Inorg. Biochem.* **5**, 1–38.
2. Munro, H. N. & Linder, M. C. (1978) *Physiol. Rev.* **58**, 317–396.
3. Bomford, A., Conlon-Hollingshead, C. & Munro, H. N. (1981) *J. Biol. Chem.* **256**, 948–955.
4. Brown, A. J. P., Leibold, E. A. & Munro, H. N. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1265–1269.
5. Costanzo, F., Santoro, C., Colantuoni, V., Bensi, G., Raugei, G., Romano, V. & Cortese, R. (1984) *EMBO J.* **3**, 23–27.
6. Costanzo, F., Colombo, M., Staempfli, S., Santoro, C., Marone, M., Frank, R., Delius, H. & Cortese, R. (1986) *Nucleic Acids Res.* **14**, 721–736.
7. Hentze, M. W., Keim, S., Papadopoulos, P., O'Brien, S., Modi, W., Drysdale, J. Leonard, W. J., Harford, J. B. & Klausner, R. D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7226–7230.
8. Santoro, C., Marone, M., Ferrone, M., Costanzo, F., Colombo, M., Minganti, C., Cortese, R. & Silengo, L. (1986) *Nucleic Acids Res.* **14**, 2863–2876.
9. Leibold, E. A. & Munro, H. N. (1987) *J. Biol. Chem.* **262**, 7335–7341.
10. Cairo, G., Bardella, L., Schiaffonati, L., Arosio, P., Levi, S. & Bernelli-Zazzera, A. (1985) *Biochem. Biophys. Res. Commun.* **133**, 314–321.
11. Chou, C. C., Gatti, R. A., Fuller, M. L., Concannon, P., Wong, A., Chada, S., Davis, R. C. & Salser, W. A. (1986) *Mol. Cell. Biol.* **6**, 566–573.
12. Aziz, N. & Munro, H. N. (1986) *Nucleic Acids Res.* **14**, 915–927.
13. Rogers, J. & Munro, H. N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2277–2281.
14. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
15. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
16. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
17. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
18. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
19. Schwarzbauer, J. E., Tamkun, J. W., Lemischka, I. R. & Hynes, R. O. (1983) *Cell* **35**, 421–431.
20. Tamkun, J. W., Schwarzbauer, J. E. & Hynes, R. O. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5140–5144.
21. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
22. Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhrone, L. & Treiber, C. (1970) *Virology* **40**, 734–744.
23. Arber, W., Enquist, L., Hohn, B., Murray, N. E. & Murray, K. (1983) *Lambda II* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) 433–466.
24. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
25. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
26. Leibold, E. A., Aziz, N., Brown, A. J. P. & Munro, H. N. (1984) *J. Biol. Chem.* **259**, 4327–4334.
27. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
28. Boyd, D., Vecoli, C., Belcher, D. M., Jain, S. W. & Drysdale, J. W. (1985) *J. Biol. Chem.* **260**, 11755–11761.
29. Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276.
30. Ford, G. C., Harrison, P. M., Rice, D. W., Smith, J. M. A., Treffry, A., White, J. C. & Yariv, J. (1984) *Philos. Trans. R. Soc. London Ser. B* **304**, 551–565.
31. Dorner, M. H., Salfeld, J., Will, H., Leibold, E. A., Vass, J. K. & Munro, H. N. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3139–3143.
32. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
33. Stevens, P. W., Dodgson, J. B. & Engel, J. D. (1987) *Mol. Cell. Biol.* **7**, 1751–1758.
34. Disbury, J. R., Theil, E. C., Kaufman, R. E. & Dickey, L. F. (1986) *J. Biol. Chem.* **261**, 949–955.
35. Shull, G. E. & Theil, E. C. (1982) *J. Biol. Chem.* **257**, 14187–14191.
36. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
37. Zucker, M. & Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.
38. Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2850–2854.
39. Zahringer, J., Baliga, B. S. & Munro, H. N. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 857–861.
40. Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. (1986) *Science* **234**, 47–52.